The Dynamics of Obsolescence and the Challenges of Legacy Data Storage

Gordon E. MacKinnon
DataSure Services
P.O. Box 42016, 2200 Oak Bay Avenue
Victoria, BC CANADA
V8R 6T4
E-mail: mackinnon@datasure.com
Tel: +1-250-598-6831
Fax: +1-250-598-6841

Abstract

The constant improvement in mass storage technology is not without risks. The dynamic nature of the storage technology and market demands for improvements in price and performance put at risk the long-term viability of data stored on legacy media.

The risks come from a number of factors at work in the market and industry. Despite the efforts of groups such as AIIM, IEEE and THIC toward the adoption of open standards there is still a tendency of manufacturers and developers to create proprietary solutions, ostensibly to enhance performance through unique advances.

The problems arise on a number of levels; in media, hardware, software and operating system to varying degrees depending on the media, form factor, format and maturity of the technology. One of the few constants is that the market itself is by far the most significant influence. What we will cover in this paper are specific examples of these kinds of problems, and some suggested solutions.

Issues

Our examination will cover two areas. We will look at the comparative performance of some of the storage options and their strengths and weaknesses. We can then develop strategies to minimize exposure to obsolescence, and minimize the impact of technological change to best maintain the data and its accessibility.

In order to examine the problems we must start with some of the variables of long term data storage choices currently available in the hardware / media. There are five major competing objectives: access time, throughput, capacity, reliability and price. These are competing in as much as the two main technologies, tape and removable disks have specific advantages and disadvantages in these areas. The solution of best fit for the capacity and performance may not provide the best value in the longer term.

Access Time

The random access performance of the optical, magneto-optical CD-R, and robotics with multiple drives can improve access time.

- Kodak ODW25 14” WORM 25GB - 700ms
- Kodak OD2000 series14” WORM jukebox - 6.5 second Disk exchange
- Magneto-Optical 2.6 GB - 19ms access time

- Pioneer DRM1004X CD-R Jukebox - 110ms
- DLT7000 - 60 second average access time per tape
- Advanced Metal Evaporated Tape / AIT - 28 second average access time per tape

Throughput

Similarly throughput can also be addressed through arrays with multiple drives to multiply throughput. There are two measurements of throughput, burst and sustained transfer rates, and we will look at sustained transfer rates for single drives.

- SONY AIT - 6 MB/sec
- DLT 7000 - 5 MB/sec
- DDS-3 - 1.2 MB/sec
- Magneto-Optical 2.6GB - 6 MB/sec
- Pioneer DRM1004X 12x CD-R Jukebox - 1.8 MB/sec

Capacity

The most obvious issue regarding mass storage is capacity. This can be addressed in two ways, through capacity of the individual media pieces, and capacity of robotic arrays. Individual drive capacities are are shown below.

- SONY DIR-1000 96 GB - 82 GB formatted
- 14” Optical WORM from Kodak provides - 25 GB per disk
- DLT7000 - 35 GB
- SONY AIT SDX-300C - 25GB
- DDS-3 - 12GB
- Magneto-Optical - 2.6GB, 2 sides
- Magneto-Optical - 5.2GB 2 sides
- CD-R - 650 MB

In both areas, tape typically has the advantage.

Reliability

Regarding longevity, disk media comes out ahead in reliability, with a lifespan estimated in the range of 30 to 100 years. Older tape media suffers from degradation much earlier.

- Advanced Metal Evaporated Tape (SONY AIT, EXABYTE Mammoth) - 30 years
- DLT - 30 years
- Magneto-Optical - 30 years
- CD-R - 100 years

The real cost of reliability is hidden in the maintenance costs as the data is migrated more frequently. This includes the cost of additional media over the life cycle of the data and the cost of additional resources required for migration. Depending on the amount of data and the frequency of migration, the advantages of less expensive initial media cost could be offset by the cost of maintenance.

Price

There are three factors in the cost of technology: media costs, drive cost and less obviously, cost of ownership. In terms of cost per GB, tape has a significant lead over MO, WORM, and CD-R. Similarly, the cost of drives also favors tape. However if one is to factor in the cost of ownership, the advantage is less clear. The reason for this is that longevity of tape media is considerably less than MO, WORM, and CD-R, and as such, the issue of duplication and migration costs is introduced. Although this issue is difficult to quantify it could become a significant factor over long periods. Also influencing the cost of ownership is the physical storage space required. With the fullness of time and the cost of real estate, a premium for high storage density can be offset by lower storage space requirements and costs.

- ADIC SCALAR 458 DLT7000 48 cart 2 drive 1680 GB - $47,000, $27.97/GB
- EXABYTE EXB 480 80 cart 1 drive 1600GB - $45,000, $28.12/GB
- ADIC SCALAR 218 DLT7000 18 cart 630 GB - $20,000, $31.75/GB
- Pioneer DRM5004X CD-R jukebox, 500 disc, 325 GB - $20,000, $61.00/GB

To be considered also would be the strategy of planning for limited life cycles of systems, with migrations to new products in the not to distant future. Such a strategy presumes improved price, performance and capacity, so as to lower long term storage cost with increased density and productivity.

Critical Path and Point of Failure

Given the choice of particular technologies, we will now examine the points of failure, their causal influences and proactive measures to minimize their impact. We will look at how both physical and market issues influence the obsolescence of technology.

Physical Issues

The physical issues affecting storage have improved in recent years, as there are higher capacity devices, and many of the drives are backward compatible.

Media Degradation

This is the most obvious failure in long term storage. It can be countered by multiple archive copies, a pre-emptive migration strategy, optimum storage environments, and a choice of media influenced by the anticipated life cycle of the resident data.

Although tape is less expensive per GB, keep in mind the cost of migration and vigilance in terms of systems and personnel costs over the life cycle of the data. Modern optical and tape robotics and software can assist in the maintenance of data.

Drive Failure

As a function of their contact with media, tape drive heads are subject to replacement more frequently than the heads in optical drives. All drives are subject to wear, and maintaining critical parts inventory can assist in maintaining drives and improving drive availability in the future.

Maintaining multiple drives can mitigate this. In the case of long term archives that may not be migrated and are resident on legacy media, care must be taken to ensure that compatible and operational drives, as well as suitable systems, are available. Tests should be executed subsequent to major upgrades to ensure continued access to the data, as new systems and software may not recognize legacy drives and formats.

Operational Problems

There is the chance that through oversight or changes, legacy data is unavailable due to operational problems such as insufficient documentation and legacy procedures. Consider the consequences of password protected backup tapes without passwords or legacy applications that lack documentation. Some older systems use proprietary hardware.

Market Dynamics

Beyond the physical problems of technology, market dynamics play a big role in determining the viability of technologies. The market dynamics we will look at will examine the impact of changes in the marketplace that influence the critical areas of media, hardware, archival software, and operating systems. Over the life of the data, the changes induced by the marketplace are most likely to influence the availability and prevalence of various technologies. The fast pace of change and innovation combined with the vagaries

Media

As storage demands grow we find capacity demands continue to increase. One of the consequences of the growth of capacity is the increasingly shorter production cycle for media of lesser capacity. This is not of itself a crucial problem because of the usual backward compatibility of newer drives and software. Aggressive price performance demands of the newer technology limit compatibility to not so distant generations. The limitations of past technologies may be surpassed by the demands of the marketplace, causing a migration to new form factors and formats. This can lead to a proliferation of legacy media form factor and formats.

This problem can be compounded by reorganization within firms and agencies where legacy formats and systems are orphaned by standardization, leaving media without the systems that created them.

Hardware / Software / Operating Systems

In the case of manufacturers and developers the issues of profitability and market-share/mind-share in relation to the industry as a whole are most critical. The first concern regarding profitability is of course the long-term survival of the manufacturers and developers, and their continuing ability to support and maintain migration paths and future iterations of technology. Their failure, poor performance, or perceived poor performance can have irrevocable effects on the market share and mind-share of the firms. With a lower profile in the marketplace, the technologies of the affected firms are less likely to attract broad support and compatibility from other industry players, and less support further

marginalizes the technologies. Poor profit performance of some products can lead to refocusing of corporate resources

With the increasing scrutiny of the corporate world by investment fund managers, the fiscal performance of the individual firms is examined in great detail on a quarterly basis. The consequences of meeting or not meeting performance expectations can have a large impact on share values. This puts pressure on firms to deliver high profit margins.

Poor fiscal performance can lead to rumors of takeovers and uncertainty, resulting in erosion of confidence in the market, the marketplace, and sometimes the loss of valuable human resources within the organization. Pressures of the financial market can make firms attractive takeover targets, with the inevitable result that competing product lines are rationalized.

Dramatic response to poor financial performance through restructuring can lead to the loss of key personnel and morale, short-term market share attrition, which is balanced against the benefit of longer term stability.

Technological obsolescence can happen quickly as new storage requirements surpass capacities. This is most true with nonstandard technology. Without aggressive marketing and production to gain marketshare and create an installed base early, nonstandard technology often remains a niche market until superseded by capacity demands.

Production

Within manufacturers, products compete for limited production facilities: the premature discontinuation of older technology facing intense competition and falling profit margins, allowing them to concentrate resources on newer and more profitable technologies. This in itself is not as much of a short term risk because the newer products are usually backwards compatible. In the long term, however, this leads to shorter product cycles and the proliferation of legacy equipment for the consumers. Low inventory production methods and shortages of key components from time to time cause some products to be unavailable in mid production cycle. This can cause problems for clients by causing them to wait or choose a competing technology and increase proliferation if the need is more immediate.

Similarly the increased scrutiny also causes software development firms to focus their critical human resources on the largest market segments, ignoring or delaying the development of products and drivers of the marginalized technology, perpetuating the cycle of decline of these products. We are led to examine the corporate structure of the manufacturers and developers and its role in the marketplace.

Niche Software

These are companies that specialize in specific market segments. Their small areas of expertise promote broad support of hardware within their area of specialization. They add to the variety of solutions available for the end user. With mergers and acquisitions, some of the smaller firms are vulnerable to larger rivals, or often the intellectual properties of some firms are of enough value to warrant a takeover. The purchasing firm saves time and development costs, and is given an instant entry into new market segments, or expanded product lines.

The challenge of supporting a range of underlying OSs and diverse automated hardware brings different approaches to file storage, introducing the issue of file storage management systems. This eventually creates a problem in migration to other file formats, and the

potential for orphaned legacy file systems, no longer supported, or no longer supporting critical hardware.

Strategies

Although there are no perfect solutions to these issues, there are strategies that can help to limit the impact of the risks.

Considering the longer term costs of ownership, the premium price of AIT, DLT 7000 or 5.2 GB M/O drives are more than offset by the longevity and density of the media and the youth of the technology. Depending on your speed and access demands, higher capacity solutions deliver better value in the long term.

One of the big concerns would be for lessening importance of the long-term interests of the storage users in the decision processes of the manufacturers and developers. The larger financial and corporate interests of the manufacturers and developers ultimately determine further development or discontinuation of technology.

The promotion and wide spread adoption of open standards lessens the risk of being limited by proprietary technologies.

The adoption of technology standards within organizations can lead to more homogenous technological environment. The homogenous environment offers the advantage of hardware redundancies and greater flexibility to restructure within the organization. The standardization occurs at the expense of users' ability to access specific non-standard technology that may be better suited to specific tasks, and the individual requirements are compromised for the standards.

At the File Storage Management System level, there have been first some efforts to adapt a standard tape format to include File Level Metadata in a tape format [1]. This has evolved into the work of File-Level Metadata for Portability of Sequential Storage (FMP) Study Group [2], which has culminated in the Association for Information and Image Management International ANSI/AIIM MS66 Proposed standard, Metadata for Interchange of Files on Sequential Storage Media Between File Storage Management Systems (FSMS) [3].

To this end public and private sector customers can specify that proposals require adherence to these standards. Early support from large clients in this manner can help to establish market momentum and an installed base for the standards, reinforcing the commitment of vendors to the standards, and attracting new participants. This in turn creates a broader installed base and helps to build momentum.

Continuous monitoring and migration can help to ensure the viability of legacy data, and its compatibility with current technology.

The maintenance of hardware, software, O/S and operations manual archives can help to ensure the accessibility of obsolete technologies; this is best complemented with ongoing testing to ensure the viability of legacy data.

Conclusion

With an explosion of on-line data driven by low cost disk storage, capacity and performance demands for removable media will grow and find new technologies.

Manufacturers and developers will evolve with the market place, and technologies will be superseded.

Most of all, continuous vigilance of the technologies and industry is essential to best deliver proactive responses to inevitable change, and the long-term commitment of appropriate budget resources to accomplish these ends. As a user of the technology you are not just buying the technology but also a migration path to the future.

References

[1] A Straw Man Proposal for a Standard Tape Format, Ben Kobler, Goddard Space Flight Center, and Joel Williams, Systems Engineering and Security.
http://ses2.ses-inc.com/~joelw/TAPE_FMT/Straw_Man.html

[2] Progress in Defining a Standard for File-Level Metadata, Joel Williams, Systems Engineering and Security.
http://ses2.ses-inc.com/~joelw/TAPE_FMT/Format_Progress.html

[3] Metadata for Interchange of Files on Sequential Storage Media Between File Storage Management Systems (FSMS) Association for Information and Image Management International ANSI/AIIM MS66 Proposed.
http://ses2.ses-inc.com/~joelw/TAPE_FMT/RSD011098.html